DOCUMENT RESUME

ED 252 564 TM 850 033

AUTHOR Penfield, Douglas A.; And Others
TITLE Analysis of Key Education Instrumentation.
PUB DATE Aug 84
NOTE 17p.; Paper presented at the Annual Meeting of the American Psychological Association (92nd, Toronto, Ontario, Canada, August 24-28, 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Auditory Perception; *Culture Fair Tests; Evaluation Criteria; Nonverbal Tests; Occupational Tests; Predictive Measurement; Visual Perception; *Vocational Aptitude; *Vocational Evaluation
IDENTIFIERS *Key Assessment System; Test Batteries

ABSTRACT

The Key Assessment System, consisting of test instruments which measure psychological functioning, work related competencies, and attitudinal and motivational characteristics, is described. The system is a vocational assessment battery designed to differentiate levels of psychophysical capabilities in a nondiscriminatory manner. It provides a mechanism for identifying vocational capabilities and making optimal vocational recommendations. The Key Tabletop Lab and the Key Audiometer are designed to minimize the threatening aspects of testing. Computer technology is utilized to produce and maintain norms specific to the sample being tested. When large groups of individuals from distinct populations are tested, it will be possible to develop unique norms for each group. Separate evaluative criteria will be available for distinct populations. Theory suggests that the more specific the criteria, the more precise the recommendation. To make the system useful in the prediction of performance in practical areas, a core set of physiological functions was established to identify performance capabilities based upon human attributes. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. A. Penfield

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

[X] This document has been reproduced as received from the person or organization originating it.
[ ] Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED252564

ANALYSIS OF KEY EDUCATION INSTRUMENTATION

Douglas A. Penfield Alvin Krass Roseann M. Conlon

KEY EDUCATION INC.

INTRODUCTION

Assessment of human functioning arose from the scientific endeavor to understand more about the nature of performance. The earliest explanations of human functioning evolved in the late 1800's when scientists measured simple human psychofunctioning in the hope of understanding issues related to motivation and individual differences. Through extensive testing and experimentation, performance standards were established for many areas of neuropsychological functioning of such facilities as vision, electrical activity of the brain, cardiac rate, and respiratory rate.

By 1900, although performance standards had been established, individual capabilities were assessed through clinical observations. Pragmatic efforts to utilize and understand the various elements of human performance led to the identification of specific abilities which were fundamental requirements for effective performance. Because educational and institutional placements were made at this time on the basis of clinical observation only, a more consistent method of classifying behavioral capabilities was sought.

In response to this need Alfred Binet developed a measure which he called an Intelligence Quotient (Terman & Merrill, 1960). Age-based standards were developed for normal performance of specific activities. Binet's underlying assumption was that the rate at which an individual had learned in the past was a reliable predictor of future performance and that skills accrued with age. His work was the first attempt to establish a performance standard, or norm, which related to

TM 820 033

Presented at the 92nd Annual Convention of the American Psychological Association at Toronto, Ontario, Canada, August 1984.

something called 'intelligence'. Binet's measurement approach, in which current achievement is compared with age-based standards, remains prevalent in psychological assessment today.

The massive testing programs conducted under the auspices of the Armed Forces during World War I presented researchers with the opportunity to identify and measure many factors associated with psychophysical functioning. Psychological researchers learned that many variables influence performance, including physical well-being, motivation, personality traits, prior experience, acquired knowledge and skills, as well as inherent ability. While numerous tests have demonstrated their usefulness for predicting future performance, there is uncertainty as to whether many of these tests are measuring basic capacity for learning.

THE KEY SYSTEM - EARLY OBSERVATIONS

Early work with atypical populations demonstrated to Key personnel that available standardized measures used for the assessment of intellect had, in many instances, little predictive validity. Children with language problems and learning handicaps, as well as those who had grown up in families with little intellectual stimulation, were often labeled as impaired even though empirical observations indicated that their functional capabilities on many tasks were normal.

In an attempt to generate an assessment system which could predict performance in practical areas, work functions were analyzed in order to develop measures which would identify performance-based attributes, rather that relying solely on performance in academic or verbal skills. A core set of physiological functions was established to identify performance capabilities based upon human attributes. Among the basic competencies identified were hearing and visual acuity, visual and auditory memory, visual and auditory processing, intersensory processing of

stimuli, manipulation and recreation of items from memory, manual dexterity, general physical well-being, persistence, speed and effectiveness of non-verbal learning, and the ability to generalize learned phenomenon through a process of mental abstraction. These were the essential psychophysical issues on which The Key System is based.

During the conceptual period, a number of researchers working in areas related to the interests of Key personnel were identified. Alexander Luria (Gardner, 1980), the Russian psychologist, in his work with mentally impaired subjects had observed that individuals who were brain damaged were capable of effectively mimicking a demonstrated task, but they were not capable of extending what was observed so as to be able to apply their knowledge in the performance of different, though similar, tasks. Normal individuals with advanced learning capabilities, however, controlled their behavior by observing, analyzing and abstracting truths observed in one context and then applying them to another. Cattell (1971), another learning theorist, introduced the notion of the crystallized and fluid aspects of learning ability. He postulated that tests of abstract reasoning such as series problems, classifications and figural analogies measure the fluid aspect of learning ability, while tests of vocabulary, reading comprehension, and verbal skills measure the crystallized component.

The goal of The Key System was to determine 'tasks' which would reliably differentiate levels of psychophysical capabilities in a nondiscriminatory manner. Evidence seems to indicate that the verbal component of standardized tests was frequently biasing results so that an attempt was made to develop a test system that would be nondiscriminatory. It was necessary, therefore, to develop a set of tasks which would be minimally language dependent or non-verbal and which would be cross-cultural in application. These tasks must also produce objective,

directly observable and quantifiable results, as opposed to data which needed to be clinically interpreted and was, therefore, subject to rater bias and inconsistency.

INITIAL DEVELOPMENT

In his work in schools and as a consulting psychologist, The Key System's originator, Alvin Krass, Ph.D., worked extensively with the educationally handicapped, a population of youngsters who were tested repeatedly throughout their school years. Because most tests were presented in paper and pencil format, Krass observed that the test mode itself contained an inherent motivational factor. In the assessment system which he conceived, Krass integrated applied performance tasks which utilized modern technology and departed from the conventional 'nuts and bolts' work sample format, as well as the paper and pencil test mode.

The test apparatus designed by Krass was specifically constructed taking into consideration those human factors which enter into the test environment. The Key Tabletop Lab and the Key Audiometer were built to be 'user friendly' so as to minimize the threatening aspects of testing which contribute to subject anxiety and diminish individual performance. Sketches of these instruments are shown in Figure 1 and Figure 2. The Key tests were also designed to be independent of experience factors. In this way, the tests were not biased in favor of those individuals who come from more privileged backgrounds or who have the benefit of varied life experiences. Studies thus far have supported this premise (Penfield & Conlon, 1983).

NORMS

Central to the concept of nondiscriminatory test practice is the process wherein individual performance is rated by comparison to some metric or performance standard. This interpretive procedure is the norming concept. Standardized

tests establish norms based on the performance of a 'standardization sample'. This group norm serves as the basis for rating all individuals who subsequently take the test. From a measurement perspective, this format often presents technical difficulties for the user who must select the standardization sample with which his local sample is to be compared. In many instances this interpretive design is both cumbersome and predictively ineffectual.

The most effective standard for interpreting individual performance is the local norm. That is, the metric produced by the group to which the individual belongs. The Key System utilizes computer technology to produce and maintain norms specific to the sample being tested. No attempt has been made to develop standardized national norms. Instead norms are generated in an ongoing manner, with each individual contributing to the standard against which performance is evaluated.

DESCRIPTION OF THE KEY SYSTEM

Tests performed within The Key System are designed to tap basic elements of psychophysical functioning. That is, those underlying perceptual skills which are fundamental for learning. Psychophysical is a term often used synonymously in the literature with psychophysiological, physiological, psychomotor or neuropsychological. As the terminology itself suggests, this area of study is concerned with complex human behaviors. Specifically, with those behaviors or activities in which stimuli are perceived and sensation or response is evoked. Within this stimulus-response context, both the central nervous system and the autonomic nervous system are involved and the research interest is focused primarily on the explanation of psychophysical functioning as it relates to more complex processes (Hirschman, Melamed, & Oliver, 1982), such as attention, arousal, orientation, emotion and learning.

The Key System subscribes to the theory that the effectiveness of job performance largely depends upon the compatibility of the match between the person and the job requirements. Many factors, however, are involved in on-the-job performance, and in order to determine an optimal placement, all of these indicators should be assessed. The work-related functions tested within the Key Assessment System are listed in Table 1. They are accompanied by their vocational corollaries and a listing of the specific test used to evaluate each function.

Developed as a vocational assessment battery appropriate for a wide range of populations, The Key System was designed to produce comprehensive results in a time-efficient and cost-effective manner. The system provides a nondiscriminatory mechanism for identifying vocational capabilities and making optimal vocational recommendations. Selected tests are integrated into the assessment process so that individual performance can be evaluated from a multi-faceted perspective. The system combines historical, demographic, functional and motivational data with measures of motor skills and work-related competencies and interests using a variety of test modes which can be categorized as follows:

Measures of Psychophysical Functioning

(a) Hand Strength and Manual Persistence
(b) Visual and Auditory Reaction Time
(c) Combined Visual-Auditory Reaction Time
(d) Visual and Color Acuity
(e) Hearing Screening, Auditory Localization, and Auditory Memory
(f) Visual Memory
(g) Fine and Gross Motor Skills

Measures of Work-Related Competencies

(h) Literacy
(i) Reading Ability
(j) Computational Ability
(k) Spatial Relations
(l) Non-Verbal Learning Ability

Attitudinal and Motivational Measures

(m) Supervisory Knowledge
(n) Expressed Vocational Interests
(o) Personality Factors
(p) Personal, Historical and Demographic Data/Structured Interview

Individual performance on all measures included in The Key System, even standardized tests published by other companies, is rated by comparing the individual test results with continuously updated norms derived from the group to which the individual belongs (i.e. individuals with shared characteristics - sex, age, level of education, occupation).

Reliability Study

No assessment of a test battery would be complete without a thorough study of the reliability of the instrumentation. Many components of The Key System have been subjected to a rigorous review of their reliability. Some of the instruments have already been standardized so that norms and psychometric characteristics are readily available in the literature. For locally constructed instruments, data were collected on samples of normal adults. Test-retest, split-half, Kuder-Richardson 20, and/or Cronbach's Coefficient Alpha reliabiliity coefficients (Thorndike & Hagen, 1969) have been computed. For a split-half form of reliability, the reliability coefficient was adjusted using the Spearman -Brown Prophecy Formula. A summary of the findings giving the (a) test (b) type of reliability, (c) sample size, and (d) reliability coefficient(s) is presented. In some instances a description of the process for consolidating data is also outlined.

Test	Type of Reliability	Sample Size	Reliability
Hand Strength and Manual Persistence	Test-Retest	195	Right Hand = 0.91 Left Hand = 0.92

The five right-hand grip readings obtained in the TEST portion were summed and correlated with the five right-hand grip readings obtained in the RETEST. The process was repeated for the left-hand grip readings.

Test	Type of Reliability	Sample Size	Reliability
Reaction Time to Visual, Aduitory, and Visual-Auditory Signals	Split-Half	118	Light = 0.80 Sound = 0.82 Light + Sound = 0.83
	Cronbach's Alpha	118	Light = 0.70 Sound = 0.77 Light + Sound = 0.80

Initial examination of the raw data indicated a great deal of variation in the second trial of each Reaction Time Subtest (series of five trials). This variation was interpreted as an attentional lapse and a decision was made to disregard the responses made on the second trial of each subtest. The remaining trials (1, 3, 4, & 5) were divided into two groups with the response to Trial 1 assigned to Group I and the response to Trial 3 assigned to Group II. The responses to Trials 4 and 5 were randomly assigned to either Groups I or II on the basis of a coin flip. If a 'head' appeared, Trial 4 went into Group II and Trial 5 into Group I. The two trial scores for each subject in Groups I and II were then added together to form a composite score.

Test	Type of Reliability	Sample Size	Reliability
Non-Verbal Learning Ability	Split-Half	200	0.93
	Cronbach's Alpha	200	0.91

Trial 1 performance time was omitted since performance time on this trial depended largely on chance. Performance times (Trials 2-5) for a sample of 200 normal adults were randomly divided into two groups. The response for Trial 2 was placed into Group I, while the response for Trial 3 was placed into Group II. A coin was flipped to determine where scores from Trials 4 and 5 would be assigned. If a 'head' appeared, Trial 4 went into Group I and Trial 5 into Group II. If a 'tail' came up on the flip, the assignments were reversed. The two trial scores for each subject in Groups I and II were then added together to form a composite score.

Test	Type of Reliability	Sample Size	Reliability
Fine Motor Skills	Split-Half	104	0.63
What's Next	Kuder-Richardson 20	60	0.77
Auditory Acuity	Test-Retest	53	Low Vol, Low Freq = 1.00 Low Vol, Med Freq = 1.00 Low Vol, High Freq= 0.83 Mid Vol, Med Freq = 1.00
Auditory Memory	Test-Retest	53	Mid Vol, Med Freq = 0.81 SLLL Mid Vol, Med Freq = 0.70 SSLL
Arithmetic Skills	Split-Half	89	0.90
	Kuder-Richardson 20	89	0.87

For the instruments presently analyzed, the reliabilities ranged from a low of 0.63 for the measurement of fine motor skills to a high of 1.00 for assessing auditory acuity. As additional data become available, reliability information will be updated.

Test Bias

To assess the potential for bias in Key instrumentation, test results from a group of correction officers were examined to determine whether data showed a tendency to produce race and sex differences. The sample consisted of 216 adults ranging in age from 18 to 59. There were 202 males nd 14 females. One hundred sixty seven applicants were Caucasian and 49 were Black. Variables were identified as being continuous or categorical. For the continuous variables (i.e. level of education, hand strength, reaction time, etc.), Caucasians and Blacks as well as Males and Females were compared using t-test procedures. For the categorical variables (i.e. learning code, auditory acuity, visual acuity, etc.), comparisons between groups were carried out using a chi-square analysis.

An instrument will be viewed as unbiased with respect to race if the scores for Blacks and Caucasians are not statistically different from each other. A similar interpretation will hold for the sex variable. If differences are found between the two races and the two sexes, the difference could be due to test bias or it could be due to actual differences in performance between the two groups.

Only two variables produced statistical differences between Caucasians and Blacks. The mean scores on Learning Code - Demo Trial Time [Caucasian Mean = 7.19; Black Mean = 11.36] and Learning Code - Average Independent Time [Caucasian Mean = 6.82; Black Mean = 8.45] were significantly higher for Blacks than for Caucasians. Relative to possible sex differences, three variables were significant. Females had a higher Level of Education than Males [Female Mean = 13.21; Male Mean = 12.09]. Males scored significantly higher than Females on Average Right Hand Strength [Male Mean = 21.22; Female Mean = 15.67] and on Average Left Hand Strength [Male Mean = 20.30; Female Mean = 16.47]. This finding is not surprising since Males tend to be stronger than Females. These results are probably due more to actual sex differences than test bias. No significant sex or race differences were found for any of the categorical variables. Considering that data from seventeen variables were analyzed, there is little evidence to indicate possible sex and race bias.

Summary

The Key Assessment System is described in detail. It consists of test instruments which measure psychological functioning, work-related competencies and attitudinal and motivational characteristics. Historical perspectives were described as well as the rationale behind the test development. Special emphasis was placed on developing instrumentation that would be free of bias. There is also a description of the reliability assessment associated with many of the tests in the battery. It is anticipated in the future that Key Education will be testing large groups of individuals from distinct populations. If this does become the case, it will be possible to develop unique norms for each of these distinct populations. This means that separate evaluative criteria will be available for groups such as correction officers, park personnel, police officers, etc. The theory suggests that the more specific the criteria, the more precise the recommendation. With the additional data being collected it will also be possible to make comparisons among groups to determine whether differences in test performance do or do not exist.

REFERENCES

Cattell, R. B. Abilities: Their Structure, Growth and Action. Boston: Houghton Mifflin Co., 1971.

Gardner, Howard. The Lives of Alexander Luria. Psychology Today, June 1980, 84-90.

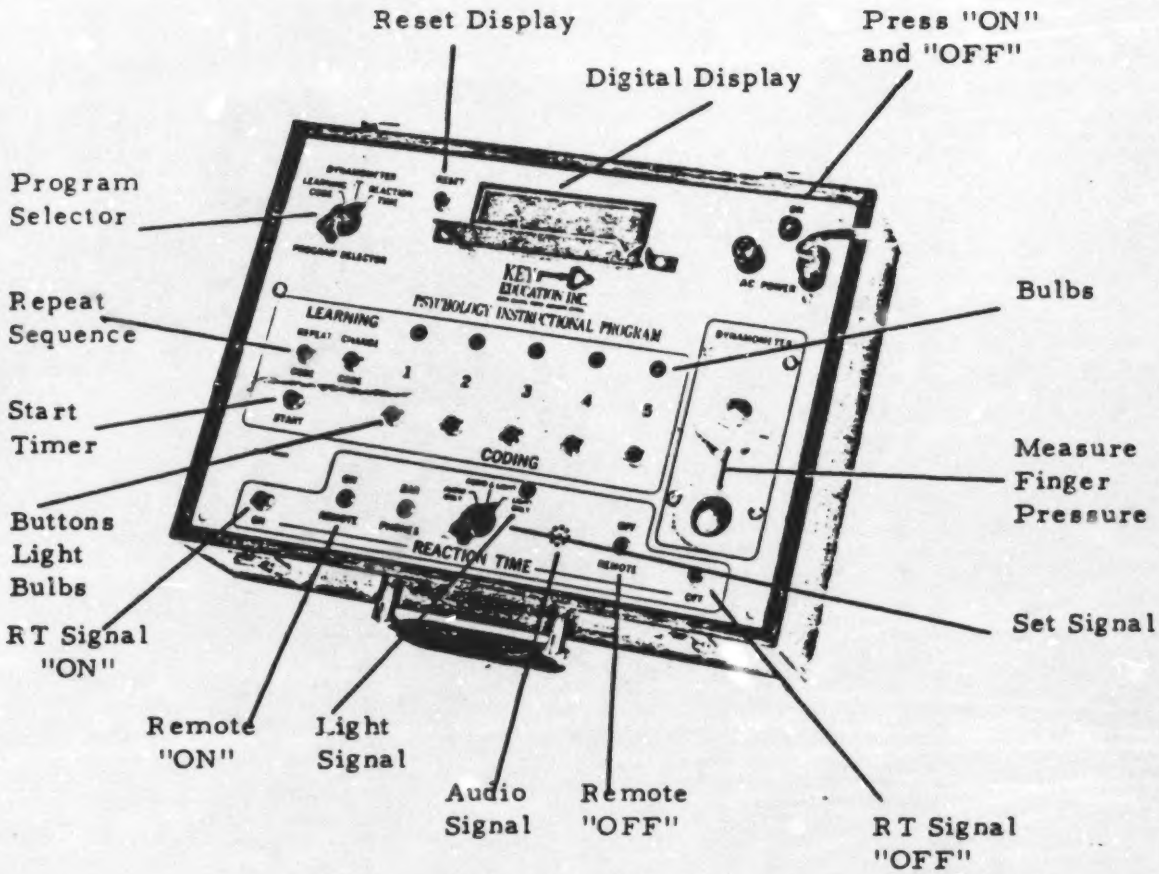
Hirschman, R., Melamed, L. E., and Oliver, C. M. The Psychophysiology of Infance. The Handbook of Developmental Psychology, 1982, 230.

Penfield, D. A. & Conlon, R. Examination of Key Instruments Test Bias. Red Bank, N.J.: Key Education Inc., 1983.

Terman, L. M. & Merrill, M. A. Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin Co., 1960.

Thorndike, R. L. & Hagen, E. Measurement and Evaluation in Psychology (3rd ed.). New York: John Wiley & Sons, 1969.

FIGURE 1



The KEY TABLETOP LABORATORY is a multi-station laboratory for assessing elements of individual functioning. The tests performed with the KEY TABLETOP LABORATORY are designed to tap basis elements of psychophysical functioning which relate to academic and vocational skills.

Equipment Specifications:

Logic and Control Circuits:	TTL Integrated Circuits
Power Requirements:	120V AC, 60Hz, 30 Watts
Portable Weight:	13 lbs.
Dimensions:	15" x 12" x 5"

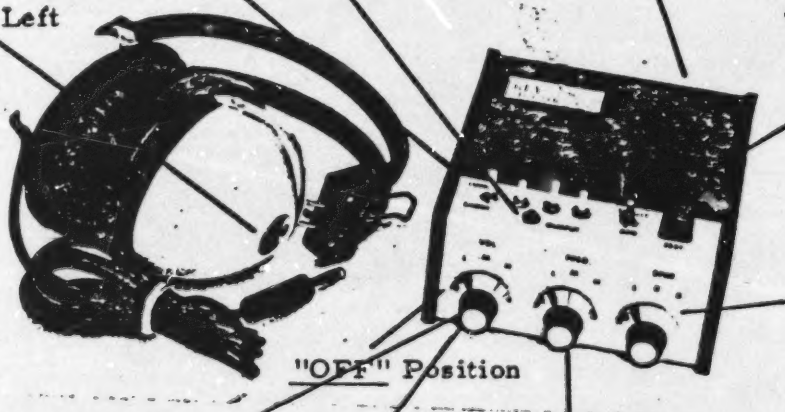
BEST COPY

FIGURE 2

With the KEY AUDIOMETER, one can perform an audiometric screening, and measure the accuracy of auditory memory. The instrument has three dials on its face. The Volume switch has an "OFF" position; turning it to any of the Volume settings automatically turns the machine "ON".

The upper switch segment has a red test signal button which turns on the signal. There is a red signal light showing that the signal is being transmitted. A green switch with two positions allows for either a single signal, or a series of signals of either Short (dots) or Long (dashes). The sequence is determined by putting the white switches either in the Long or Short position. Placing the white switches in the Long position will result in 4 Long signals, placing the switches in the Short position will result in Short signals.

The Ear Phones are labeled Left and Right and the ear phone jack fits into the back of the device.



Light for visual signal (shows where sound signal is being given).

Switch sets Long or Short signals in 4 signal series.

Speaker Jack (ear phone jack fits into socket in rear-push all the way into socket).

Speaker (Left and Right ear sides are marked).

Switch to set Single signal or 4 signal series (Single = 1 Signal) (Sequence = 4 Signals)

To the Speaker (Earphones)
L=left ear, only
R=right ear, only
B=both ears

"OFF" Position

Turn "ON" by turning switch to any Volume setting.

Frequency Settings
L=Low sounding signal (500 cycles/sec.)
M=Medium sounding signal (2000 cycles/sec.)
H=High sounding signal (4000 cycles/sec.)

To Set Volume
L=Low loudness (30 dbs)
M=Medium loudness (30-45 dbs)
H=High loudness (+55 dbs)

BEST COPY

TABLE 1

Function	Test	Vocational Corollary
Hearing	Key Audiometer	Following directions; Responding to signals and danger cues; Social interaction.
Auditory Localization	Key Audiometer	Accurately identifying the direction from which sound signals emanate.
Auditory Memory	Key Audiometer	Learning and all of its ramifications; Being able to follow directions and to remember what has been said; To retain information.
Visual Acuity	Ciba-Geigy Chart	Ability to perceive visual stimuli and perform tasks requiring good vision.
Color Acuity	Ciba-Geigy Chart	Ability to deal with color coded information.
Visual Memory	Key Audiometer & Key Tabletop Lab/ Learning Code Component	Ability to learn from and retain information gathered through sight (i.e. filling out forms, mechanical repairs).
Manual Dexterity	Key Fine Motor Skills Test	Ability to perform desterous tasks and to learn skills requiring eye-hand coordination with speed and accuracy.
Auditory Reaction Time	Key Tabletop Lab/ Reaction Time Component	Ability to respond to auditory stimuli and to perform tasks involving verbal communication.
Visual Reaction Time	Key Tabletop Lab/ Reaction Time Component	Ability to respond to visual signals
Abstract Reasoning Ability	Key What's Next	Ability to problem solve and reason in an abstract manner.
Non-Verbal Learning Ability	Key Tabletop Lab/ Learning Code Component	To be able to learn from demonstrated directions; To mimic tasks through observation and to generalize from basic task perfromance to independent function and mastery.
Response to Stress	Key Tabletop Lab/ Dynamometer Component	Speed and accuracy of hand-eye coordinated tasks when under performance stress.

TABLE 1 CONTINUED

Function	Test	Vocational Corollary
Hand Strength	Key Tabletop Lab/ Dynamometer Component	Ability to perform tasks which require hand strength; An index of general physical well-being.
Persistence	Key Tabletop Lab/ Dynamometer Component	Ability to work at repetitive tasks without significant decline in performance; A measure of motivation.
Literacy Skills	Key Form S-3	Ability to execute basic writing skills, including spelling, accuracy, quality of word usage and vocabulary comprehension.
Arithmetic Skills	Key Arithmetic Test	Ability to perform basic computational math problems accurately.
Reading Level	Wide Range Achievement Test/ Reading Scale	Ability to read at a level appropriate for vocational placement.
Vocational Interest	Key Vocational Interest Inventory & Self-Directed Search	To develop a profile of vocational interests which will indicate areas of vocational potential.
Supervisory Potential	How Supervise	Knowledge of common supervisory practices.
Personality Attributes	Sentence Completion & Graphics Projectives	Identifies personality qualities which highly correlate with management and supervisory effectiveness and indicate basic emotional stability.